Spring 2016

Prediction of Vapor/Liquid Equilibrium Behavior from Quantum Mechanical Data

Marshall J. Gnap
University of Akron, mjg78@zips.uakron.edu

Honors Research Project 2016

Prediction of Vapor/Liquid Equilibrium Behavior from Quantum Mechanical Data

Marshall Gnap

May 6, 2016

TABLE OF CONTENTS

Honors Abstract

MOSCED (Modified Separation of Cohesive Energy Density) is a particularly attractive model for activity coefficients because it offers intuitive insights into how to tune solvent-solute interactions to achieve optimized formulations. Unfortunately, only 133 compounds have been characterized with the MOSCED method. Furthermore, there is no convenient method for extending MOSCED predictions to new compounds. The hypothesis of the present research is that the surface charge density of a molecule, once normalized over the molecule surface area, provided graphically by a σ-profile from density functional theory (DFT) computations, can be used to estimate the parameters used in the MOSCED model. DFT results are readily available for 1432 compounds through a public database at Virginia Tech, and further DFT computations for new compounds are relatively quick and simple due to minimal additional molecular properties.

The predictive functions were regressed based on 4375 binary solution infinite dilution coefficients. The average logarithmic deviation for the predictive MOSCED model was 0.280 while using the original correlative model had a deviation of 0.106 compared to 0.183 for the UNIFAC model. Phase equilibrium predictions were also compared where various models were used for interpolating finite compositions. The average percent deviations of the pressure for the 39 binary systems tested were 17.39% for Wilson, 18.90% for NRTL, and 13.83% for SSCED.

Executive Summary

Commonly used activity models such as UNIFAC and MOSCED (Modified Separation of Cohesive Energy Density) rely on empirically characterized parameters to predict phase behavior for petrochemical mixtures. MOSCED is particularly attractive in many situations because it offers intuitive insights into how to tune solvent-solute interactions to achieve optimized formulations. Unfortunately, only 133 compounds have been characterized with the MOSCED method, compared to UNIFAC which has been developed using thousands of compounds. Furthermore, extending UNIFAC to new compounds is straightforward through the group contribution concept, whereas MOSCED simply requires more experimental data specific to the new compounds.

The following paper details a sufficiently simple method to calculate MOSCED parameters in order to determine infinite dilution activity coefficients based on density functional theory (DFT) calculations provided by Virginia Tech. The hypothesis of the present research presents that the surface charge density of a molecule, provided by a σ-profile from the DFT computation, can be used to estimate the α, β, and τ parameters used in the MOSCED model. By defining a charge density threshold for regions of hydrogen bonding, the probability of the charge exceeding that threshold and size of the surface can be correlated to determine the parameters for the molecules. By assuming the remaining surface charge potential that does not contribute to hydrogen bonding represents the polarity of the molecule, the polarizability parameter τ can also be determined.

The current number of characterized compounds for the MOSCED model is 133, of which only 89 are considered to have non zero acidity and basicity parameters. This number limits the possible use of the method due to the low number of molecules exhibiting interactions other than dispersion forces. The proposed method allows for the expansion of the model to 1432 different compounds along with any molecules that are characterized in the future by a σ-profile. The correlated functions were regressed based on 4375 binary solution infinite dilution coefficients provided by Lazzaroni et al. and the deviation from the experimental data was calculated based on the logarithmic ratio of the calculated versus experimental [(4)]. The resulting average deviation for the MOSCED model with correlated parameters was found to be 0.28 while using the original parameters tuned to the experimental data had a deviation of 0.106. The UNIFAC model was also compared for binary solutions in which the functional groups were defined and the resulting deviation was found to be 0.183. The calculated infinite dilution coefficients were then used to interpolate the entire phase behavior of a binary system across the composition range. The "simplified MOSCED" (SSCED), Wilson, and NRTL models were chosen to test the accuracy of the method based on the low number of parameters needed to define the interaction energies for the system. The resulting phase equilibrium predictions were compared to experimentally determined results from the Danner and Gess database [(10)]. The average percent deviations of the pressure for the 39 binary systems tested were 17.39% for Wilson, 18.90% for NRTL, and 13.83% for SSCED.

Overall the method was able to determine the infinite dilution coefficients for the binary solutions with reasonable accuracy. Major deviations from the experimental value could be seen as the coefficient increased beyond 10^2 indicating poor accuracy for highly positive non-ideal

systems. The phase envelopes had large variance in the accuracy between models and interaction types. The original development of each of the models had different intentions and system types in mind, so the accuracy of the results depends highly on prior knowledge of the difference between models. The original implementation of MOSCED also failed to include amines and contained a small number of carboxylic acids. These systems in particular experienced large deviations in comparison to experimental data for all three models used. Due to the main goal of simplicity, more complex equations and adding a large number of functional-group specific parameters were not considered in the analysis. The accuracy of the infinite dilution coefficients and phase envelopes could be increased significantly with future work by utilizing more complex equations and regression. Larger data sets with a greater variety of binary solutions would also allow for more accurate predictions. In order to properly predict amines and carboxylic acids, parameters could be retrofitted based on phase behavior data. The final recommendation is the extension of coefficients to other interpolating models with greater complexity for the phase behavior. With the described limitations, the described method is able to provide a large degree of thermodynamic insight in equilibrium systems for broad combinations of molecules. For a design engineer the interaction type and degree of deviation from ideality can be quickly determined in lieu of charts or graphs when considering solvents for process design.

The project was unique due to the fact that the analysis performed had not yet been completed by other researchers. Attempting to characterize parameters, instead of directly calculating activity coefficients, based on surface charge density has not been the subject of other research papers. Over the course of the experience, I was able to gain technical skills in molecular modeling, programming, literature research, time management, and description of equilibrium systems. The project personally increased my creativity and critical thinking skill in respect to coming up with and justify solutions to problems that do not have a definitive answer. I was given a large amount of independence that made me push myself to meet deadlines increasing my time management skills as well.

1. Introduction/Background

A large number of chemical processes involve the purification of mixtures in industry such as liquid extraction, distillation, and crystallization. Any physical separation of components relies on the characterization of the phase equilibrium associated with the chemicals involved and a large quantity of studies have been dedicated trying to find a sufficient method of predicting the interactions involved with vapor-liquid equilibrium (VLE), liquid-liquid equilibrium (LLE), and solid liquid equilibrium (SLE). Obtaining experimental data for the exact mixture of components with the correct equilibrium temperature and pressure is difficult due to the scattered amount of information and expense in acquiring the reports necessary, if they are even available. When designing a process with physical separation unit operations in mind, experimental data is not always available and the equipment to produce the information most likely will not be readily on hand, so predictive models that can provide an accurate interactive system are very useful. Frank et al. discuss the use of a priori methods of prediction and their ability to sufficiently rank credible solvents for practical engineering use in industry [(1)].

Infinite dilution coefficients can be used to characterize the vapor-liquid, solid-liquid, and liquid-liquid phase behavior that underlies these unit operations. By utilizing Modified Raoult's Law, the desired phase composition, pressure, or temperature needed can be determined provided the activity coefficient used was determined by a model that can accurately predict the system and the required parameters needed are known for the components used. For purification of mixtures, the largest deviation from ideality occurs in the dilute solution system [(2)]. The ability to predict infinite dilution coefficients based on simple guidelines of hydrogen bonding could provide valuable design formulations and useful thermodynamic insights.

One of the most common methods of phase equilibria prediction for multicomponent and binary solutions is the UNIFAC method which estimates the activity coefficient based on individual group contributions [(2)]. The method requires experimental data of mixture and knowledge of the structure of the molecules to properly characterize the parameter for the contributing energies and their interaction with other groups. The method remains limited due to its lack of explicit representation of specific interactions occurring and in situations where a particular group has not been characterized.

A second accepted method of predicting the phase equilibria of a solution is based on modifications to the solubility parameter to characterize individual interactions from intermolecular forces. Traditionally the regular solution theory (RST) and its extension, the Hansen model, have been used in this category [(2)]. The RST model fails to account for negative deviations, where the activity coefficient is less than one, indicating an affinity of the molecules caused by hydrogen bonding or polarity inherent in the structure which limits its use in systems with high polarity molecules.

The Modified Separation of Cohesive Energy Density (MOSCED) model attempts to account for the deviations by separating the energy into five parameters responsible for the individual interactions [(3)]. Additionally, the specific molar volume of the liquid, V_i, for each component is used to account size differences of the molecules as well as the temperature of the

system, T(K), to find the infinite dilution activity coefficient γ_i^{∞}. Eqns. 1.1-1.7 show the MOSCED equation for determining the infinite dilution coefficient of a binary mixture.

$$ln\gamma_2^{\infty} = \frac{V_2}{RT}\left[(\lambda_2 - \lambda_1)^2 + q_1^2 q_2^2 \frac{(\tau_2^T - \tau_1^T)^2}{\psi_1} + \frac{(\alpha_2^T - \alpha_1^T)(\beta_2^T - \beta_1^T)}{\xi_1}\right] + d_{12} \quad (1.1)$$

$$d_{12} = 1 - \left(\frac{V_2}{V_1}\right)^{aa} + aaln\left(\frac{V_2}{V_1}\right) \quad (1.2)$$

$$aa = 0.953 - 0.002314((\tau_2^T)^2 + \alpha_2^T \beta_2^T) \quad (1.3)$$

$$\alpha_i^T = \alpha_i\left(\frac{293}{T(K)}\right)^{0.8}; \beta_i^T = \beta_i\left(\frac{293}{T(K)}\right)^{0.8}; \tau_i^T = \tau_i\left(\frac{293}{T(K)}\right)^{0.4} \quad (1.4)$$

$$\psi_1 = POL + 0.002629\alpha_1^T \beta_1^T \quad (1.5)$$

$$\xi_1 = 0.68(POL - 1) + [3.24 - 2.4\exp(-0.002687(\alpha_1\beta_1)^{1.5})]^{\left(\frac{293}{T}\right)^2} \quad (1.6)$$

$$POL = 1 + 1.15q_1^4[1 - exp(-0.002337(\tau_1^T)^3)] \quad (1.7)$$

Lazzaroni et at. later revised the initial interaction parameters determined for the first 89 compounds: λ dispersion parameter, τ polarity parameter, α hydrogen-bond acidity parameter, β basicity hydrogen-bond parameter, and q a factor ranging between 0.9-1 using 133 solvents with 6441 experimentally determined binary solution infinite dilution activity coefficients and tested the method to evaluate solid-liquid equilibria [4]. The parameters were regressed by minimizing the function in Eqn. 1.8 which represents the error.

$$Error = (\ln\gamma_{calc}^{\infty} - ln\gamma_{exp}^{\infty})^2 \quad (1.8)$$

This objective function has the advantage of matching the % deviation when deviations are small, but is unbiased by large negative deviations. For example, a calculated value that is too high by a factor of two would indicate 100% deviation whereas a value that is too low by a factor of two would be only -50% deviation. The logarithmic objective of Eq. 1.8 gives symmetric values of ±0.69 for either of these cases.

Other such models include the Flory-Huggins, Wilson, and Non-Random Two Liquid (NRTL) models which have roots in the RST model. All the modified models have interaction parameters specific to the compounds in the system that are fitted based on experimental data. The additional characterization of compounds can increase the utility of the MOSCED method as an easily calculable method to determine relatively accurate phase equilibrium in the dilute system. The usefulness of the method could be increased if more parameters could be characterized without the use of experimental data. Additionally, infinite dilution coefficients determined by the MOSCED model could be used to quickly characterize interaction parameters in other simpler models to develop entire phase envelope information for binary systems. The following paper proposes a method to determine MOSCED pure compound parameters based on sigma profiles calculated by Mullins et al. in order to extend the available number of defined compounds from 133 to 1432 molecules[5]. Mullins et al. have calculated σ-profiles and any future molecules which will be described by a sigma profile [5]. The method can provide equally

or greater amount of understanding and information for a system when considering a solvent or separation scheme than conventional charts, with no additional work.

2. Project Premises and Methods

Quantum mechanics calculations can be used to gain insight into charge density on the surface of a molecule in order to determine the degree of interaction associated with contact between molecules. The Conductor-like Screening Model for Real Solvents (COSMOS-RS) developed by Klamt and Eckert utilizes σ-profiles calculated using a density functional theory (DFT) in order to determine two dimensional electron density profiles for a given molecule by determining the charge density over the surface[6]. For the graphical representation, σ refers to the surface charge density in units of ($C/Å^2$) and p(σ) ($Å^2/(C/Å^2)$) is the probability density of the area per σ-interval having the specified surface charge[6]. The curve is normalized such that an integral of the profile would yield the total surface area of the molecule. The refinement of the parameters of the method by Klamt et al. produced various properties such as the vapor pressures of compounds and partition coefficients [7]. An activity coefficient equation for VLE is also available from Klamt and Frank utilizing the COSMO-RS in COSMOtherm implementation [8]. A similar activity coefficient model developed by Lin and Sandler that can be used for binary solutions consists of a summation over the discrete polarization intervals of the σ-profile providing a general prediction for many compounds [9]. According to the COSMO-RS model, the charge density of σ ranges over positive and negative values, where all values less than -0.0084 are attributed to acidity(α) or proton donation while values over 0.0084 are attributed to basicity(β) or proton acceptance in the traditional definition sense due to the projected field having an opposite charge of polarized section of the molecule. Figures 2.1 and 2.2 show sigma profiles provided by the Virginia Tech website for the molecules benzene and ethanol, where the vertical lines represent the cutoff threshold for hydrogen bonding [5]. From the graphical representation, benzene does not have the areas where the threshold is surpassed and therefore does not hydrogen bond as opposed to the ethanol molecule where both the positive and negative energies surpass the threshold allowing for the polar interaction. This analysis matches with conventional knowledge and expectations of the interactions of the molecules.

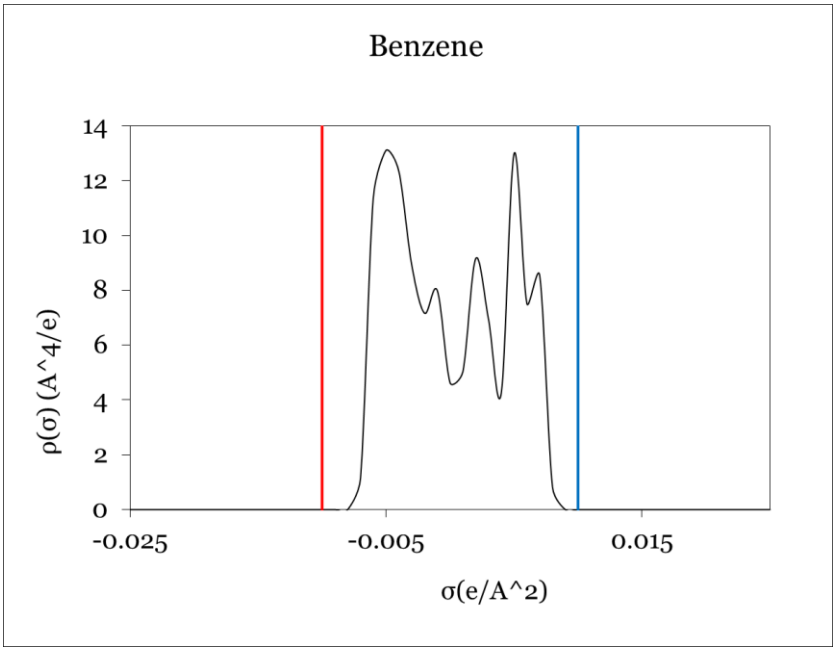

Figure 2.1. Sigma profile of Benzene molecule

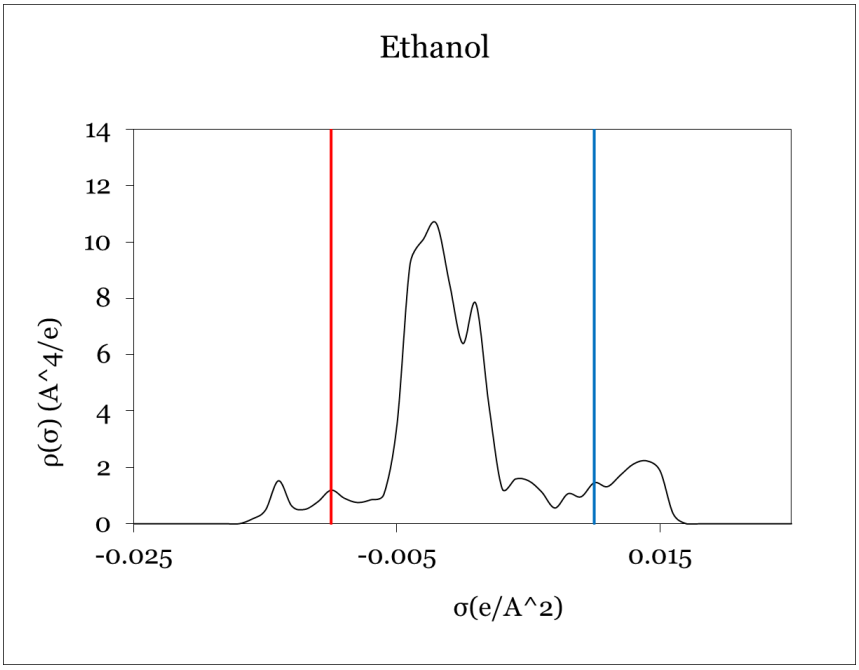

Figure 2.2. Sigma profile of Ethanol molecule

The hypothesis of the present research is that there is an analogous relationship between the surface charge density in the acidic and basic regions of the σ-profile and the α and β parameters characterized for the interaction contributions in the MOSCED model based on the intensity of the charge density in the profile for molecules that are characterized as polar and are known to have hydrogen bonding interactions.

$$\alpha \sim \int_{-0.025}^{\sigma 1} p(\sigma) d\sigma \quad (2.1)$$

$$\beta \sim \int_{\sigma 2}^{0.025} p(\sigma) d\sigma \quad (2.2)$$

For the 133 characterized solvents provided by Lazzaroni et al., corresponding sigma profiles were collected by utilizing the Virginia Tech database in order to determine the magnitude of the charge density associated with the areas crossing a defined charge density threshold, σ1 and σ2.

Initially, the charge density threshold for both the acidic and basic regions was set the same as the COSMOS-RS model, at -0.0084 and 0.0084 (C/Å^2) [(2)]. The acidic and basic parameter was then correlated to fit the parameters defined by Lazzaroni et at. in order to determine an appropriately simple fit for the calculation. Of the 133 characterized solvents, only 89 showed energy densities exceeding the designated threshold initially stated, where the additional 44 compounds were specified as non-polar and given a value of zero for the acidic and basic parameter in compliance with the Lazzaroni et al. defined values [(4)]. The remaining 89 solvents were then characterized into two groups based on whether the compound had traditional hydrogen-bonding groups in the structure to account for the high acidity or basicity parameter associated with a low integral value. The α parameter was designated and fit as a second order polynomial if the structure contained a hydroxyl group (OH=1) and a first order for the remaining solvents (OH=0). The same approach was used for the β parameter to fit molecular structures containing nitrogen based substituent groups as a second order (N=1) and the remaining as a first order (N=0). Compounds containing aromatic or halogen groups are known to experience areas of high polarity but had low integral values for the stated energy density threshold area, so two additional parameters were specified to account for structures containing these groups. Figures 2.1 and 2.2 show the resulting regressed functions of Eqs. 2.3 and 2.4 which were developed to fit the initial parameters based on the assumption of bimodal distribution of the data set. Eqns. 2.3 and 2.4 were used to calculate the two parameters as a function of the integral of the area that crossed the charge density threshold of the sigma profile. The constant parameters were later regressed based on a database of binary infinite dilution coefficients for the compounds as well as the range which the energy density threshold started at for the acidic and basic hydrogen-bonding contribution. The range allowing for the lowest deviation between the experimental values and the calculated was found to be symmetric, where σ1= -0.01 for the alpha region and σ2=0.01 for the basic region.

$$\alpha_{Calc} = \int_{-0.025}^{\sigma 1} p(\sigma) d\sigma \left[A_\alpha + OH \left(B_\alpha - A_\alpha + C_\alpha \int_{-0.025}^{\sigma 1} p(\sigma) d\sigma \right) \right] + D_\alpha + Ar_\alpha + H_\alpha \quad (2.3)$$

$$\beta_{Calc} = \int_{\sigma 2}^{0.025} p(\sigma) d\sigma \left[A_\beta + N \left(B_\beta - A_\beta + C_\beta \int_{\sigma 2}^{0.025} p(\sigma) d\sigma \right) \right] + D_\beta + Ar_\beta + H_\beta \quad (2.4)$$

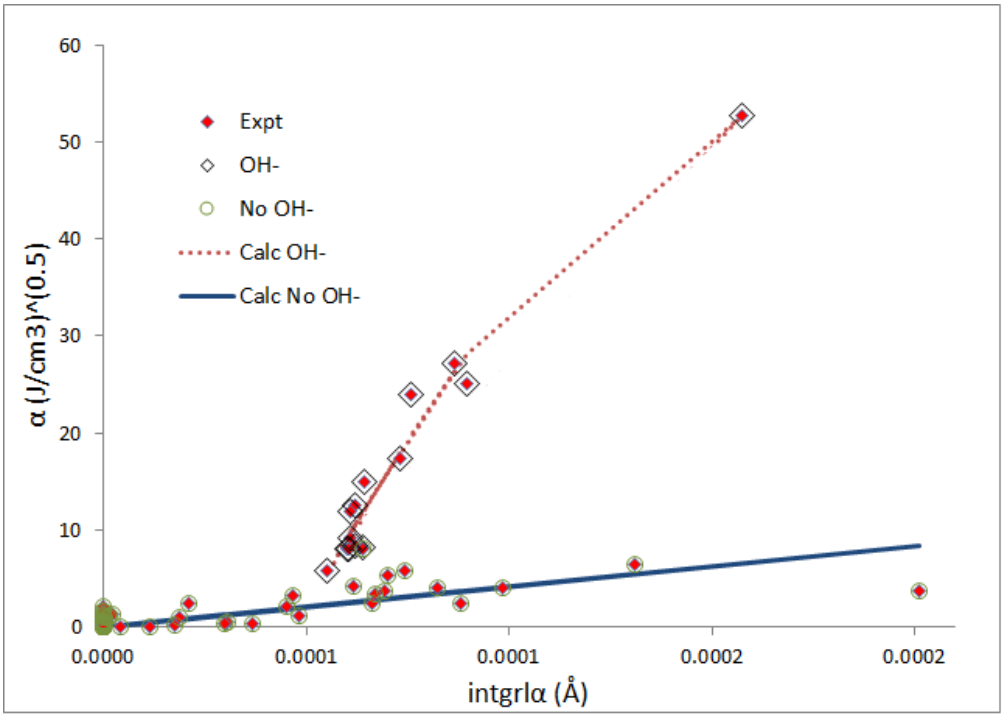

Figure 2.1. Initial correlation curve used in order to determine a suitable function for α parameter calculations. Values were regressed to match Lazzaroni et al. parameters based on a bimodal data set which was differentiated by the presence of a traditional hydrogen bonding group.

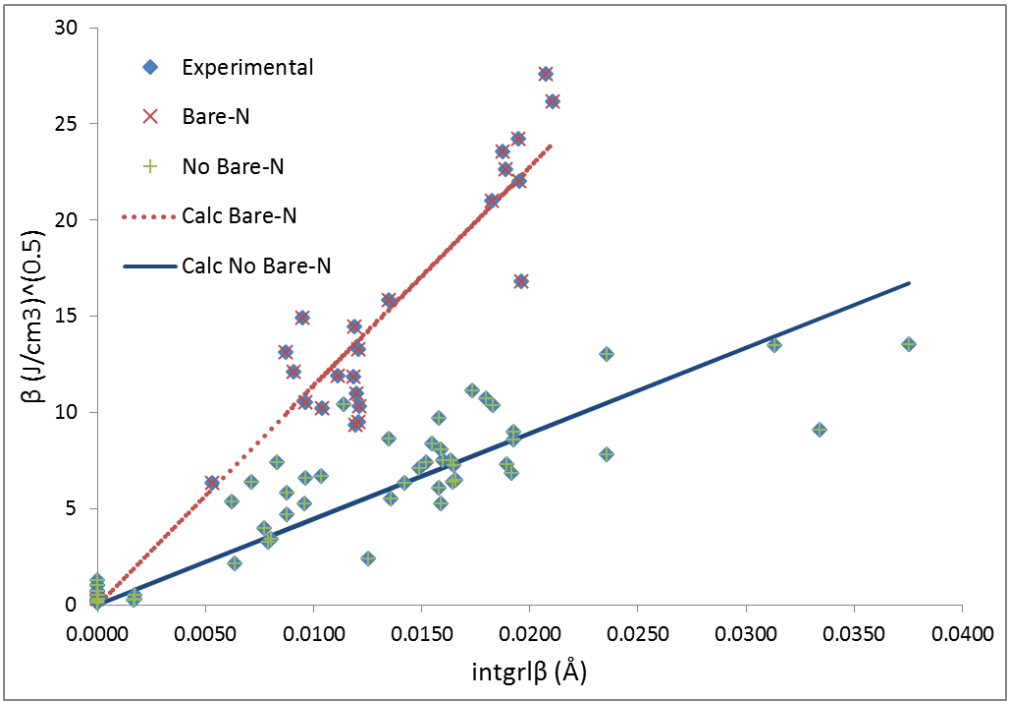

Figure 2.2. Initial correlation curve used in order to determine a suitable function for β parameter calculations. Values were regressed to match Lazzaroni et al. parameters based on a bimodal data set which was differentiated by the presence of a traditional hydrogen bonding group.

In order to calculate the infinite dilution coefficients using the MOSCED method, three additional parameters were needed to be known: the dispersion factor, polarity factor, and fitting factor *q*. For the present purpose, *q* was taken to be 1 due to a lack of direct theoretical correlation between the quantum mechanics calculations and the factor. To determine τ (polarity factor), a weighted function was developed in order to provide greater emphasis on charge density further from the surface of the molecule designated on the σ-profile to be at point 0. The polarity contributions were to be set in between the energy threshold for the hydrogen bonding for σ1 and σ2. The function was normalized over the total volume of the molecule and fitted to a first order correlation. Figure 2.3 shows the resulting correlation of the calculated polarity factors based on Eqn. 2.5 as a function of the weighted integral function versus the Lazzaroni et al. defined values. The polarity factors were not assumed to have a symmetric distribution and were fitted directly to a first order polynomial with the major contribution factor based entirely on the negative contributions from the left half of the σ-profile.

$$\tau_{Calc} = A_\tau \left[\frac{\int_{\sigma 1}^{0} \sigma^{n_1} p(\sigma) \mathrm{d}\sigma + \int_{0}^{\sigma 2} \sigma^{n_2} p(\sigma) \mathrm{d}\sigma}{\int_{-0.025}^{0.025} p(\sigma) \mathrm{d}\sigma} \right] + B_\tau \tag{2.5}$$

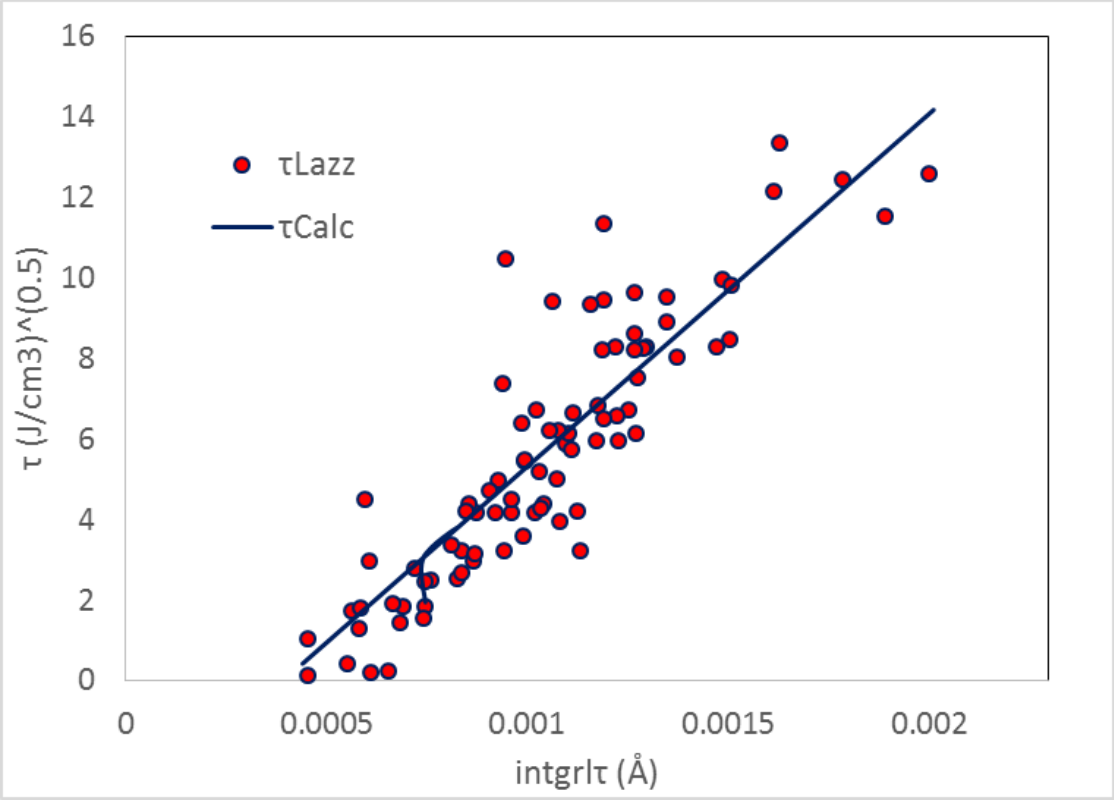

Figure 2.3. Initial correlation curve used in order to determine a suitable function for τ parameter calculations. Values were regressed to match Lazzaroni et al. parameters.

The final parameter λ (dispersion factor) was defined using the relation in Eqn. 2.6 which balances the sum of the interactive forces with literature defined solubility parameters for the compounds [(3)].

$$\delta^2 = \lambda_{Calc}^2 + \frac{\tau_{Calc}^2}{2} + S\alpha_{Calc}\beta_{Calc} \quad (2.6)$$

After the correlation of the parameters, experimental binary solution infinite dilution coefficient data where obtained for 4375 different solvent solutions and used initially by Lazzaroni et at. to regress MOSCED parameters used for the initial correlation [4]. Eqns. 2.3-2.6 were regressed using the error Eqn. 1.8 in order to match the calculated infinite dilution coefficients to the experimental data by varying the additional parameters. Additionally, the energy density threshold for the hydrogen bonding regions for the molecule were adjusted to more accurately fit the experimental data. Figure 2.4 shows a plot of the experimentally determined values versus the Lazzaroni et al. MOSCED, correlated MOSCED, and UNIFAC model. The average percent deviation in log units from the experimental data was calculated to be 0.28 for the calculated MOSCED, 0.106 for the original MOSCED parameter values, and 0.183 for the UNIFAC model.

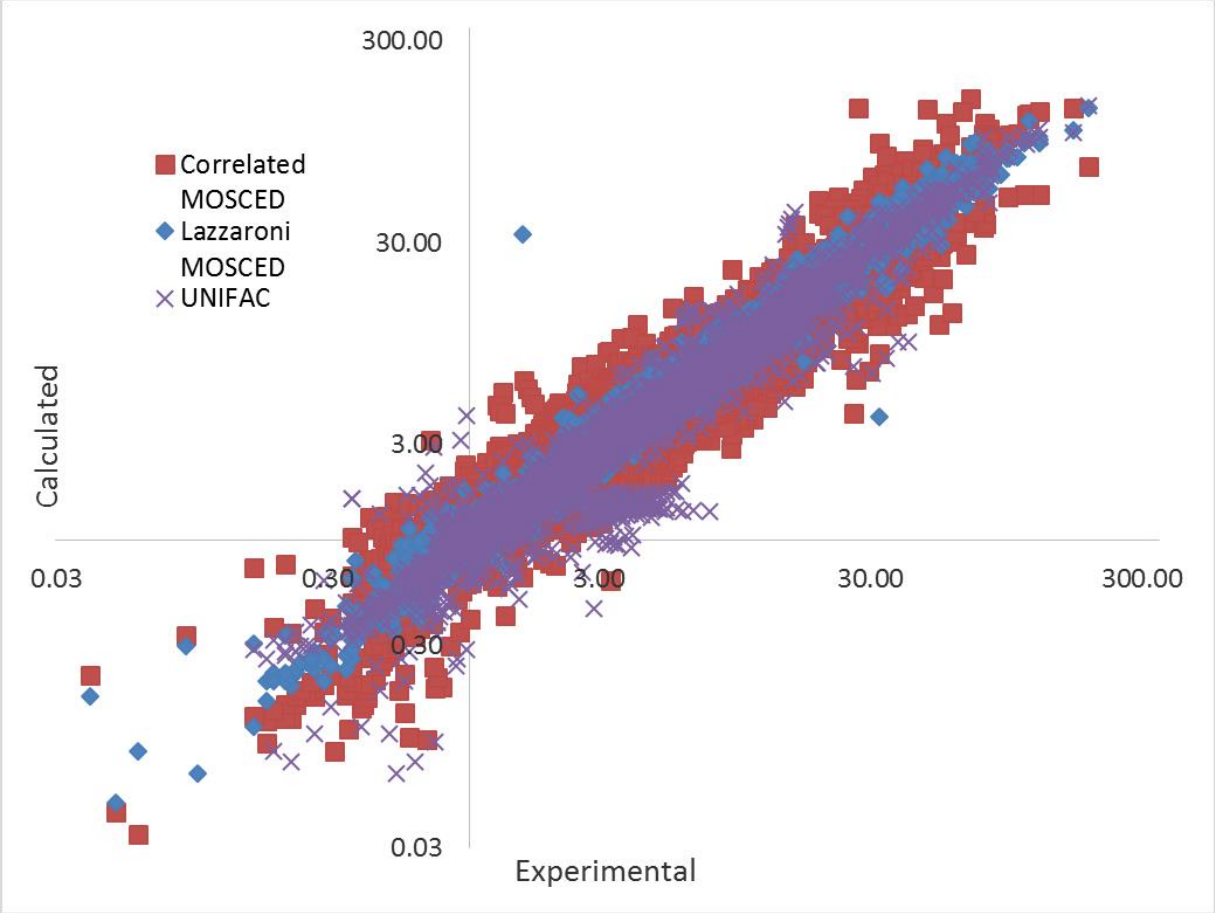

Figure 2.4. Natural log of infinite dilution activity coefficient for UNIFAC, calculated MOSCED, and original parameter MOSCED versus experimentally determined coefficients.

An additional analysis was performed to determine the accuracy of the SSCED using the calculated parameters versus MOSCED using the original parameters and UNIFAC in order to determine if the additional parameters in the MOSCED equation were necessary. The average

deviation for infinite dilution activity coefficients calculated using the SSCED model was determined to be 0.799, significantly higher than the MOSCED average deviation. Figure 2.5 shows the resulting plot comparing the three models versus the experimental data.

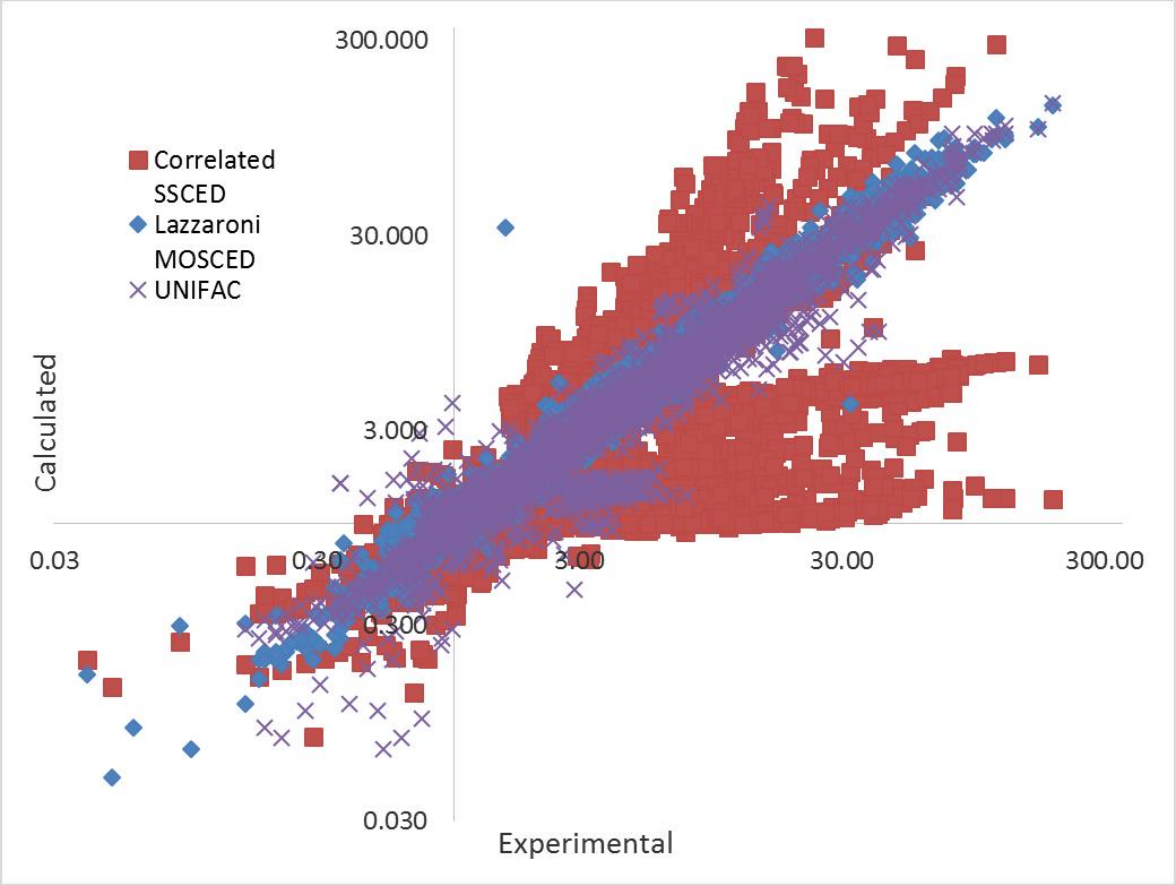

Figure 2.4. Natural log of infinite dilution activity coefficient for UNIFAC, calculated SSCED, and original parameter MOSCED versus experimentally determined coefficients.

Table 2.1 Regressed Parameters for Equations 2.3 and 2.4

	α	β
A	852.6	401.2
B	9375	938.2
C	-133167	88.13
D	-33.15	2.365
Ar	0.6703	1.055
H	0.3860	0.7067
σ	-0.01	0.01

The regressed parameters after fitting to the experimental data can be seen in Tables 2.1 and 2.2, where the sigma value represents the new range for the energy density threshold of the sigma profile for the hydrogen-bonding regions of the molecule. The constants Ar and H represent the contributions to acidity and basicity with respect to molecules that contain aromatic rings and halogens in their structure. For Table 2.2 n1 and n2 are the exponential terms in Eqn. 2.5 where the major contribution to the correlation is determined by the negative values of the sigma profile only.

Table 2.2 Regressed Parameters for Eqns. 2.5 and 2.6

	τ
A	8742.9
B	-3.430
n1	1.110
n2	29.99
S	2.071

Eqns. 2.3-2.6 were used to determine the corresponding MOSCED parameters from the σ-profiles provided in the Virginia Tech database. The resulting α, β, and τ parameters for each of the 1432 compounds can be seen in Appendix B.

Three sufficiently simple thermodynamic models were chosen to interpolate a phase envelope. Wilson's model was chosen due to the two interaction parameters being able to be calculated by the infinite dilution activity coefficients provided by the MOSCED model seen in Eqs. 2.7 and 2.8. The second model was chosen to be the NRTL, where the assumed alpha parameter for the all binary interactions was set to 0.3 and the two interaction parameters were calculated again using the activity coefficients seen in Eqs. 2.9-2.11. The final model chosen was the SSCED model which provides the total phase behavior as a simplified version of the MOSCED model [2]. The experimental data that the calculated results were compared to was primarily a database compiled by Danner and Gess to test thermodynamic models utilizing a wide variety of systems such as non-polar/polar, weakly polar/polar, and immiscible systems [10].

$$ln\gamma_1 = -ln(x_1 + x_2\Lambda_{12}) + x_2\left(\frac{\Lambda_{12}}{x_1+x_2\Lambda_{12}} - \frac{\Lambda_{21}}{x_1\Lambda_{21}+x_2}\right) \quad (2.7)$$

$$ln\gamma_1^{\infty} = -\ln(\Lambda_{12}) + (1 - \Lambda_{21}) \quad (2.8)$$

$$ln\gamma_1 = x_2^2 \left(\frac{\tau_{12} G_{12}}{(x_1 G_{12} + x_2)^2} + \tau_{21} \left(\frac{G_{21}}{x_1 + x_2 G_{21}} \right)^2 \right) \quad (2.9)$$

$$G_{ij} = exp\left(-\alpha_{ij}^{NRTL} \tau_{ij}\right) \quad (2.10)$$

$$ln\gamma_1^{\infty} = (\tau_{12} G_{12} + \tau_{21}) \quad (2.11)$$

For Eqs. 2.7 and 2.8 of the Wilson model, the x_i value represents the composition of the component, Λ_{12} and Λ_{21} the interaction parameters for the Wilson model, and γ_i the activity coefficient for the compound. Similarly, the activity coefficient, γ, for the NRTL model is also a function of the composition, x_i, as well as three separate interaction coefficients for the specific for the model: τ_{12}, τ_{21}, and α_{ij}^{NRTL}.

3. Results and Discussion

The calculated error using Eqn. 1.8 for the correlated infinite dilution coefficients calculated using the MOSCED equation was found to be roughly three times that of the Lazzaroni et al. parameters. The main source of error in the deviations from the experimental values comes from systems with large deviations from ideality in which the calculated infinite dilution coefficient was determined to be on the order of 10^2. The original parameters managed to minimize the error in cases such as a binary solution of pentane and N-methyl-formamide where the experimental coefficient was determined to be 26.4 and the correlated value was 35.5 while the original parameters produced 23.4. Smaller values for the coefficients, indicating more miscible systems, produced more consistent results, such as the case with benzene and methyl t-butyl ether where the experimental was 0.91, the correlated parameter 1.12, and the original parameter value 1.02. The correlated parameter, experimental compiled by Lazzaroni et al., original parameter MOSCED, and UNIFAC infinite dilution coefficient values for the corresponding binary systems can be seen in Appendix A[(4)]. The trends described above continue throughout the calculations, but provide relatively accurate predictions of the coefficient for the sign of the interaction and the magnitude of the value.

The choice of the three models used for the interpolation was based on the simplicity of the interaction parameters. The Wilson model contains only two undefined interaction parameters that can be defined based on the simplified Eqn. 2.8 which becomes a function of the infinite dilution coefficient calculated using the MOSCED model. For the NRTL model, the α^{NRTL} parameter, if unknown, can be assumed to be 0.3 making Eqn. 2.11 a function of the infinite dilution coefficient as with the Wilson model. Finally, the SSCED model is purely a function of the MOSCED parameters. A list of 39 binary solution interactions in the Danner and Gess database can be seen below in Table 3.1 with the calculated average percent deviation from experimental data.

Major deviations from experimental values could be seen in systems where there is a liquid-liquid phase split which is seen in the binary solution nitroethane+N-octane. Figure 3.1 shows the original calculated phase envelope for the three models. Major deviations from the experimental data occur in the center of the phase envelope where the vapor composition fraction for the component at the specified pressure exceeds the liquid fraction. This occurrence can be seen primarily in the NRTL model and slightly in the SSCED model. The interaction is indicative of the model attempting to predict areas of LLE. The Wilson model completely fails to predict LLE for all systems due to the nature of how the activity coefficient is calculated in the equation, and in areas where the coefficient is significantly large, the corresponding pressure for the system remains constant. For systems were the experimental data is not actually experiencing VLLE, the Wilson model performs significantly better due to its constraint than more general models like SSCED and NRTL. To accurately compare the results from the three models, areas of LLE were addressed in a manner similar to how the Wilson model accounts for the activity. If the activity coefficient were to exceed a value of ten, the pressure at that point was calculated and then averaged with vapor pressure of the pure component for the left and right side depending on the where the phase split occurred. The pressure for the system was then kept constant at this average value for the phase envelope composition values where threshold was exceeded. This modified phase envelope can be seen in Figure 3.2 and the additional graphs in Figures 3.3-3.10. The areas of predicted LLE were also taken into account when calculating the percent deviations from the experimental data in Table 3.1.

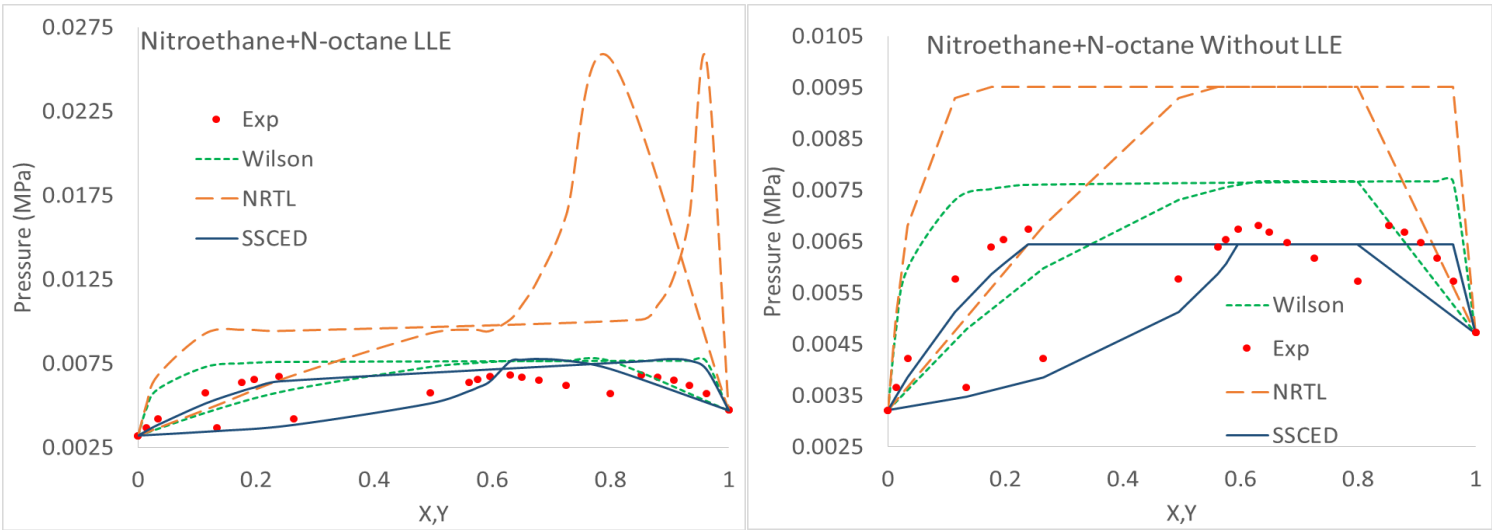

Figure 3.1 and 3.2 Vapor liquid Pxy phase diagram for the binary system Nitroethane and N-Octane designated as an immiscible system by Danner and Gess[10].

Overall the average percent deviation for the 39 solutions was found to be 17.39% for Wilson, 18.90% for NRTL, and 13.83% for SSCED. However, the variance on the percent deviations across the binary systems varies greatly due to the type of system that the model was originally developed for. From the results, the Wilson model performed, on average, better in systems where the other two models predicted LLE. However, the knowledge of areas of VLLE

is important in characterizing the phase behavior of components and the discrepancies were reduced with the addition of the modifications described above. The Wilson model also was determined to be less accurate for systems in which high polarity components interacted such as water+ethanol and n-butyl formate+formic acid. This behavior is due to the original intended design of the equation, which does not properly characterize strong polar interactions. The SSCED model had a greater accuracy for a majority of the polar interactions systems. The interaction parameter for the model is based entirely on the MOSCED coefficients and is meant to characterize polar interaction systems. For systems such as acetone+chloroform, the SSCED model resulted in 4.26% deviation versus 9.67% for NRTL and 9.60% for Wilson. Situations in which SSCED calculated more accurate results, NRTL and Wilson produced similar values. There were few binary systems where the NRTL equation resulted in a more accurate prediction than both the SSCED and Wilson model, and in those occurrences, the result was not significantly different.

Table 3.1. Average Percent Deviation for binary Systems

Binary System				Avg % Dev		
Pressure (Mpa)	Temperature (K)	Compound 1	Compound 2	Wilson	NRTL	SSCED
-	318.15	nitromethane	1,2-DICHLOROETHANE	1.93%	1.95%	0.71%
-	337.2	nitromethane	1,2-DICHLOROETHANE	1.86%	1.88%	0.80%
-	354.2	nitromethane	1,2-DICHLOROETHANE	1.71%	1.73%	0.78%
-	318.5	nitromethane	1,2-DICHLOROETHANE	1.14%	1.16%	0.07%
-	343.9	nitromethane	1,2-DICHLOROETHANE	1.03%	1.05%	0.04%
-	355.3	nitromethane	1,2-DICHLOROETHANE	1.12%	1.13%	0.20%
-	298.2	nitromethane	1,4-dioxane	0.79%	0.78%	0.04%
-	308.15	1,4-dioxane	methanol	27.03%	24.72%	22.50%
-	333.15	n-decane	acetone	12.76%	31.49%	18.63%
-	327.76	ethylbenzene	n-heptane	16.36%	17.10%	0.50%
-	328.15	1-heptene	toluene	6.39%	6.57%	1.15%
-	308.15	nitroethane	n-octane	19.18%	41.86%	5.55%
-	318.15	methanol	cyclohexane	8.07%	12.52%	55.71%
-	298.15	2-methyl-pentane	nitroethane	38.20%	46.40%	31.47%
-	298.15	n-hexane	nitroethane	34.77%	44.49%	24.06%
-	313.15	chlorobenzene	propionic-acid	38.55%	49.69%	51.05%
-	363.15	water	cyclohexanone	15.61%	11.19%	16.26%
-	317.55	water	phenol	69.72%	65.19%	13.16%
-	311.5	water	diethyl-amine	43.49%	43.39%	40.23%
-	362.98	water	pyridine	7.24%	2.73%	13.95%
-	308.15	diethyl ether	methyl-iodide	7.26%	7.74%	5.02%
-	318.15	N-PROPIONALDEHYDE	methyl-ethyl-ketone	0.75%	0.75%	0.66%
-	298.15	cyclohexane	toluene	12.66%	13.13%	5.35%
-	328.15	benzene	diethyl-amine	68.00%	68.04%	60.82%
-	308.15	acetone	chloroform	9.60%	9.67%	4.26%
-	318.15	n-butyl-amine	1-propanol	37.18%	38.36%	38.77%
-	298.15	cyclopentane	chloroform	1.10%	1.09%	3.30%
-	303.15	diethyl ether	acetone	10.81%	11.72%	1.71%
-	328.15	benzene	thiophene	1.84%	1.84%	1.67%
0.101325	-	benzene	Toluene	0.61%	0.61%	0.28%
0.101325	-	n-octane	ethylbenzene	9.97%	10.19%	1.26%
0.101325	-	dichloromethane	acetic-acid	51.33%	66.75%	60.49%
0.101325	-	n-butyl-formate	formic acid	33.05%	21.54%	22.57%
0.101325	-	formic acid	acetic-acid	3.33%	3.42%	1.36%
0.101325	-	water	methanol	7.09%	5.74%	1.08%
0.01266562	-	water	isopropyl-alcohol	43.59%	39.98%	14.21%
0.101325	-	water	ethanol	25.25%	21.87%	11.13%
0.101325	-	benzene	acetone	5.50%	5.50%	5.72%
0.101325	-	acetone	n-propyl-acetate	2.28%	2.27%	2.73%
Average				17.39%	18.90%	13.83%

Figures 3.1-3.10 show examples of interpolated phase envelopes from the database compiled by Danner and Gess[10]. Nine types of systems were designated by the database to cover the binary interactions, and a graph containing the three models interpolated using calculated infinite dilution coefficients can be seen below. For a majority of the solutions the accuracy of the model increases at the end points of the phase envelope, which indicates accurate dilute characterization based on the coefficients characterized. Excluding the interaction of N-butyl-amine and 1-Propanol, the type of deviation from ideality, positive or negative, is correctly determined by the models used in the figures below. All three models fail to characterize the system in Figure 3.3 which is assigned as a strongly polar/strongly polar interaction between the two components. The resulting graph indicates that the solution is ideal while the experimental data indicates that the binary system will result in a strong negative deviation. From the original parameter characterization, there were no amine compounds, which makes the predictions of the components in a binary solution a difficult task. For further work, the amine groups will need a special parameter based on experimental VLE data in order to reconcile the higher basicity that should be associated with such compounds. This issue can be seen again in the characterization of the parameters for carboxylic acids as seen in Figure 3.10. The high over estimation of the deviation from ideality can be attributed to an unrealistic characterization of the basicity. These parameters will need to be redefined in the future based on VLE information to more accurately predict phase behavior for anomalously basic or acidic compounds.

The graph for Figure 3.4 shows the interaction of a weakly polar/strongly polar system of components, where the SSCED model can be seen to provide the most accurate representation of the experimental data. The Wilson and NRTL models show a low degree of negative interaction for the system, which poorly characterizes the potential for polar interaction of the acetone molecule as opposed to the SCCED model. From Figure 3.6 we can see that the most accurate model is NRTL model when it comes to describing the aqueous/strongly polar system of water and pyridine. The Wilson model provides a shifted curve for the experimental data where the SSCED model seems to indicate, in the dilute water system, a LLE phase split occurring which was not completely fixed with the modifications based on the activity coefficient threshold specified at ten. The non-polar/strongly polar system shown in Figure 3.7 shows a binary system in which all three models show relatively accurate results for the system. The accuracy of the predicted infinite dilution coefficient does not seem to be an indicator of the accuracy of the interpolated phase diagram.

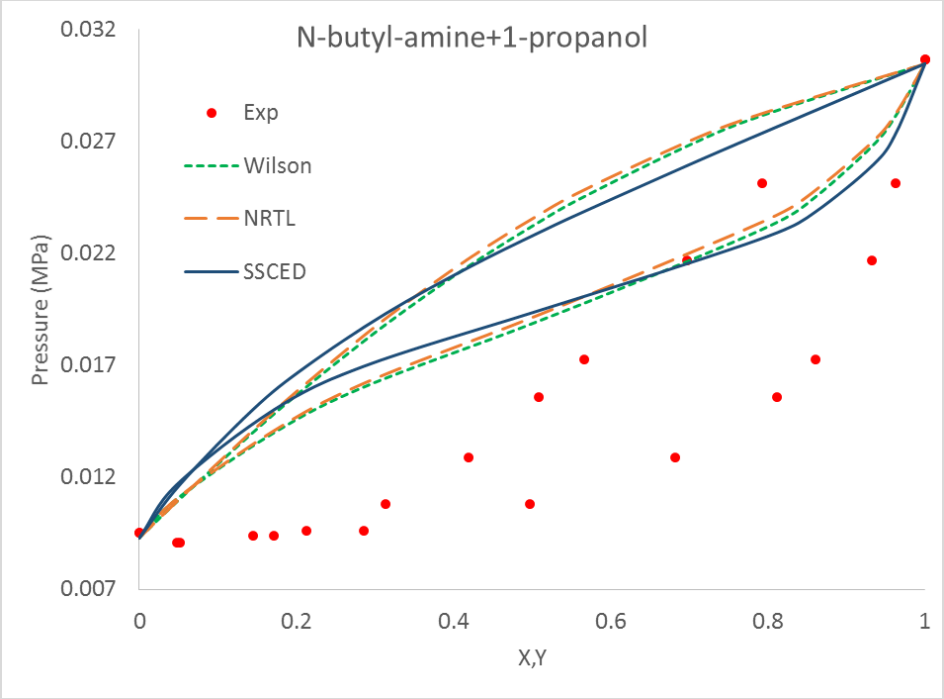

Figure 3.3 Vapor liquid Pxy phase diagram for the binary system N-Butyl-Amine and 1-Propanol at 318.15 K designated as strongly polar-strongly polar interaction by Danner and Gess[10].

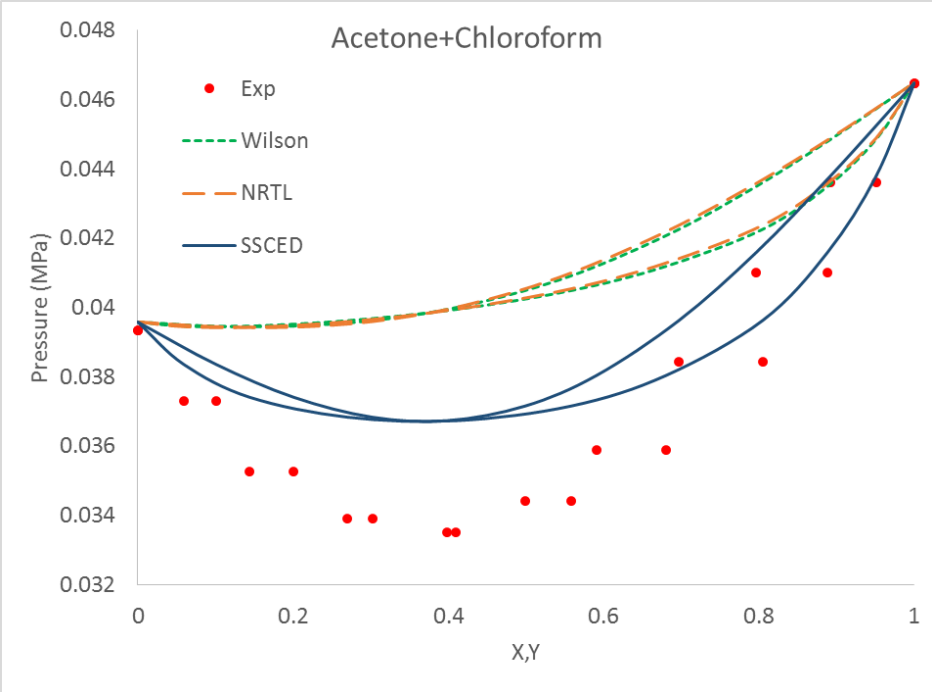

Figure 3.4 Vapor liquid Pxy phase diagram for the binary system Acetone and Chloroform at 308.15 K designated as a weakly polar-strongly polar interaction by Danner and Gess[10].

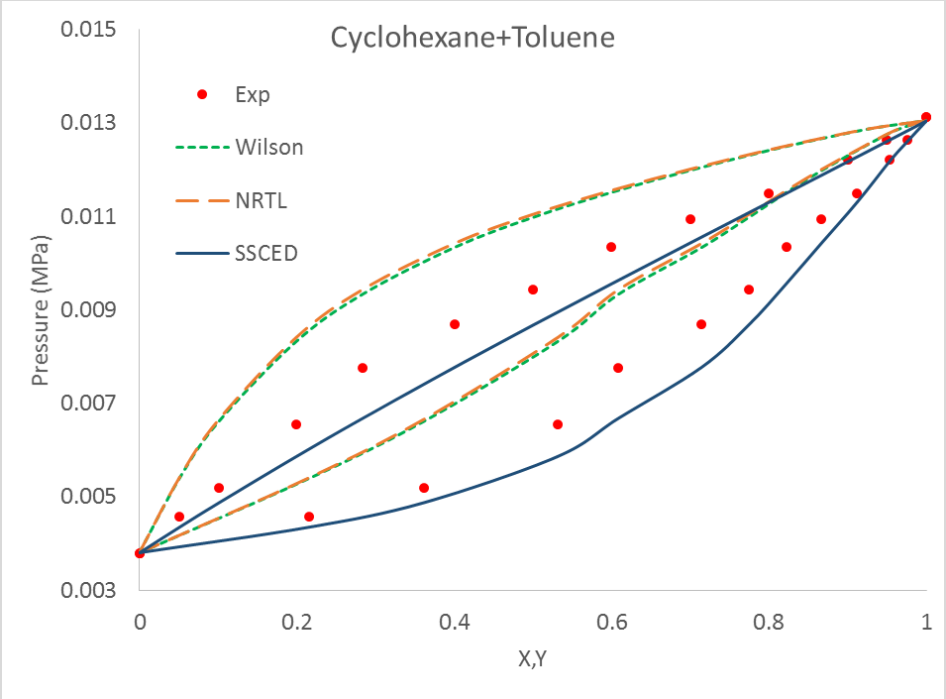

Figure 3.5 Vapor liquid Pxy phase diagram for the binary system Cyclohexane and Toluene at 298.15 K designated as a non-polar-non polar interaction by Danner and Gess[10].

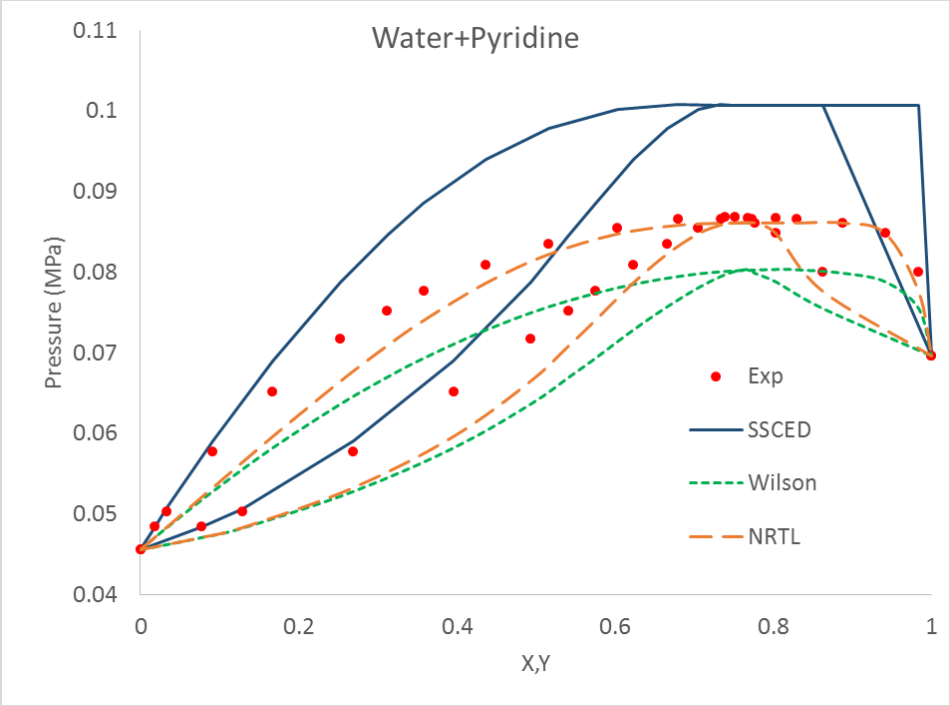

Figure 3.6 Vapor liquid Pxy phase diagram for the binary system Water and Pyridine at 362.98 K designated as an aqueous-strongly polar interaction by Danner and Gess[10].

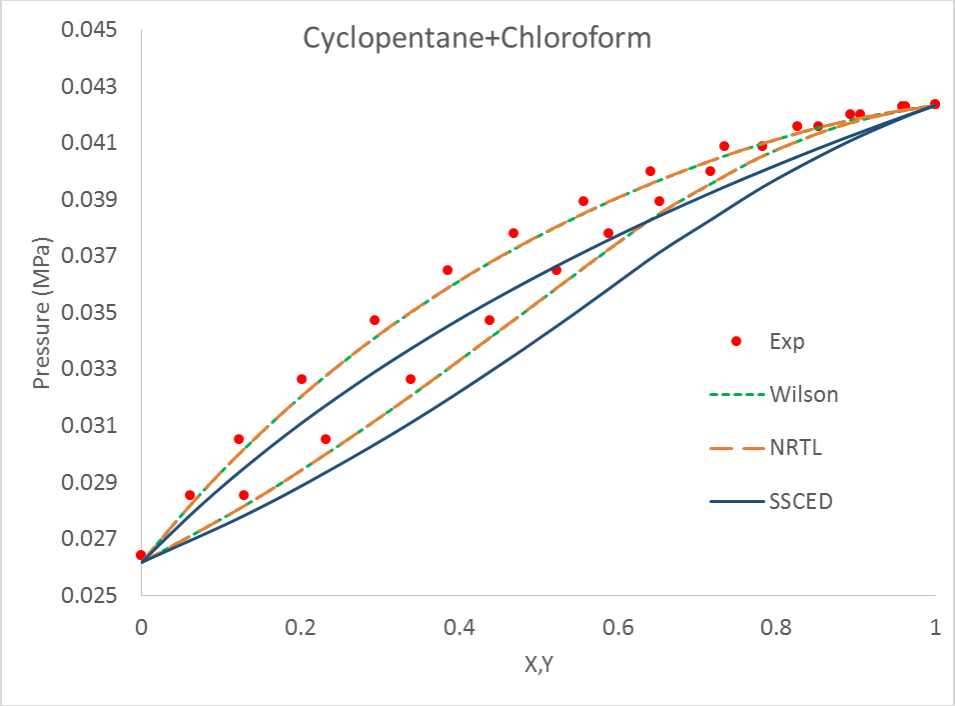

Figure 3.7 Vapor liquid Pxy phase diagram for the binary system Cyclopentane and Chloroform at 298.15 K designated as a non-polar-strongly polar interaction by Danner and Gess[10].

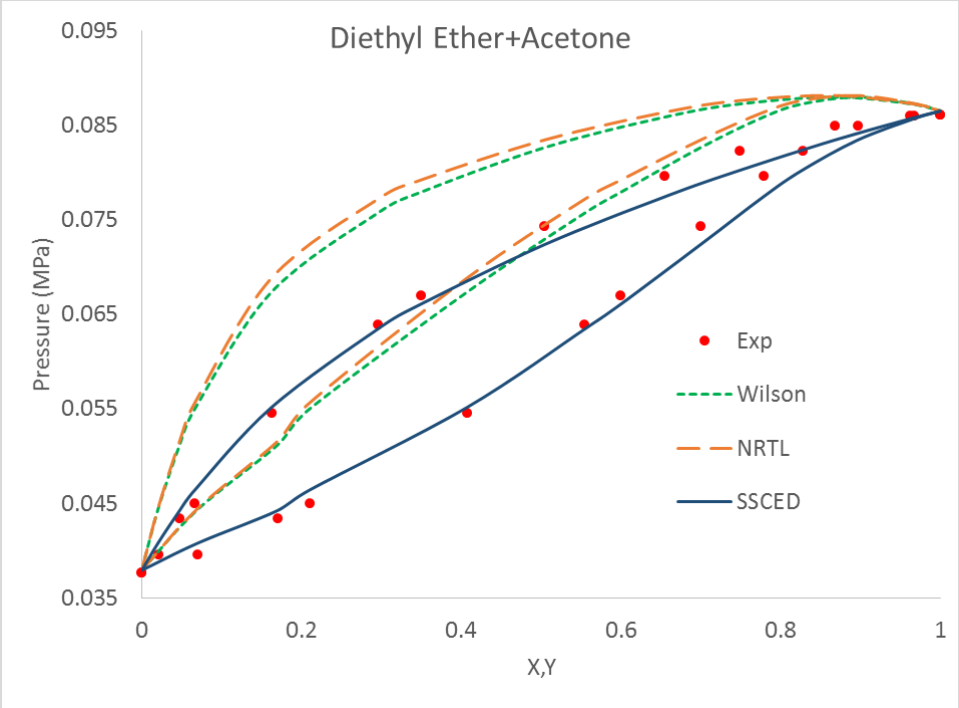

Figure 3.8 Vapor liquid Pxy phase diagram for the binary system Diethyl Ether and Acetone at 303.15 K designated as a weakly polar-weakly polar interaction by Danner and Gess[10].

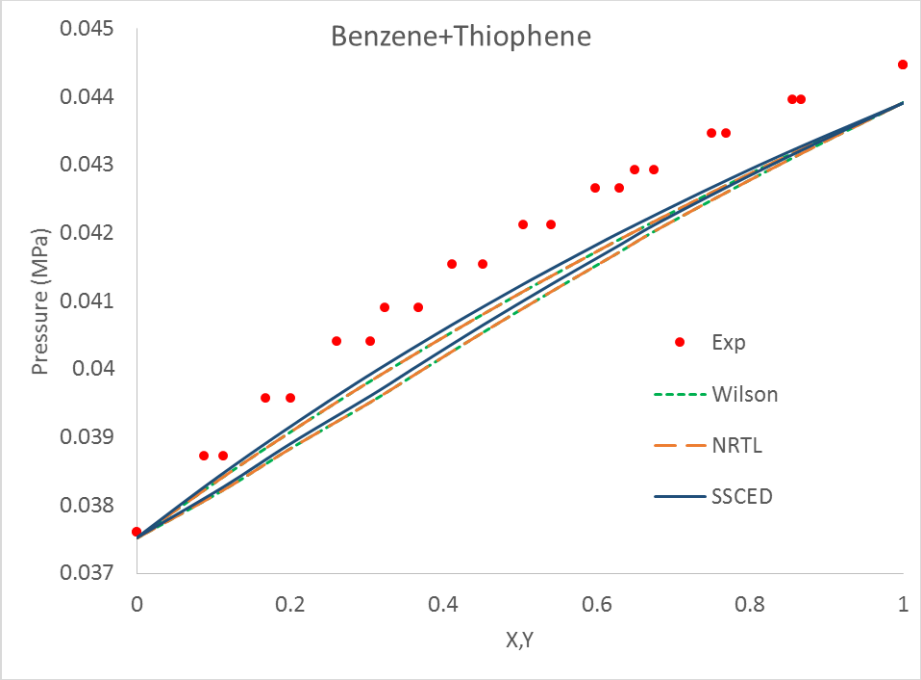

Figure 3.9 Vapor liquid Pxy phase diagram for the binary system Benzene and Thiophene at 328.15 K designated as a non-polar-weakly polar interaction by Danner and Gess[10].

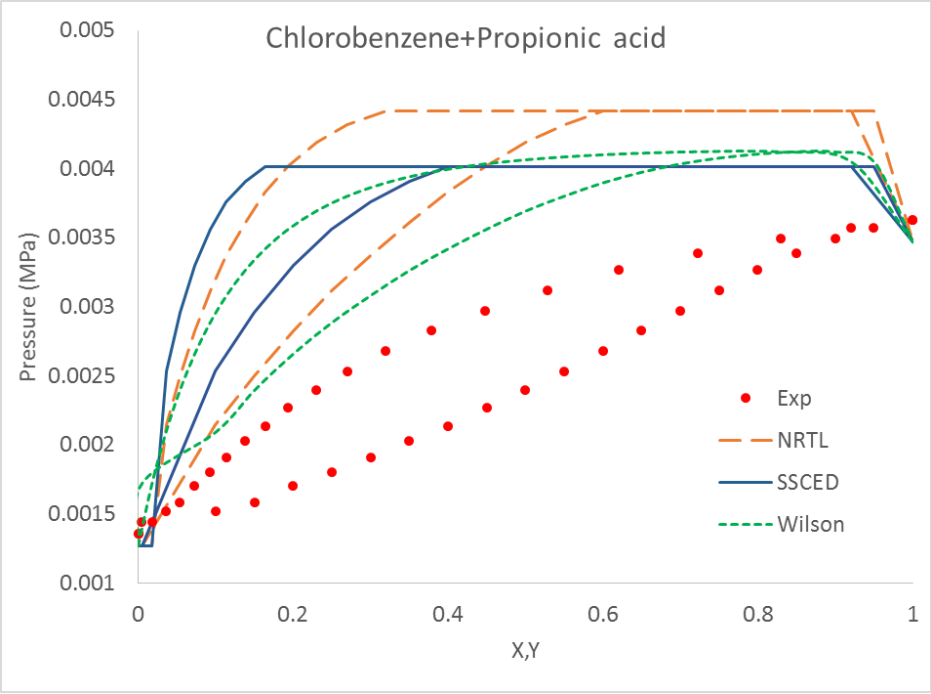

Figure 3.10 Vapor liquid Pxy phase diagram for the binary system Chlorobenzene and Propionic Acid at 313.15 K designated as a carboxylic acid system by Danner and Gess[10].

4. Conclusions and Recommendations

The overall characterization of MOSSCED parameters from σ-profiles calculated using DFT was found to have a relative accuracy for the binary solutions analyzed. When compared to the original experimental infinite dilution activity coefficient database used by Lazzaroni et al. to determine the current MOSSCED parameters, the error from experimental values was found to be roughly three times that MOSSCED model with the original parameters and twice that of the UNIFAC model. The accuracy of the method decreased significantly, in comparison with the original values, with the increased magnitude of the coefficient, but the parameters were able to still predict the positive or negative deviation and the magnitude of the coefficient relative to the experimental value. The average deviation for the correlated parameters was found to be 0.28, 0.106 for the original MOSCED, and 0.183 for the UNIFAC model.

The extension of the model to include entire VLE phase envelopes resulted in accurate predictions when both the calculated dilution coefficient was accurate in comparison to the experimental data and the model being used was appropriate for the system. For the 39 binary solutions examined from the database of Danner and Gess, the average percent deviation of the system pressure from the experimental value for the Wilson model was found to be 17.46%, 18.24 for NRTL, and 12.9% for SSCED. There existed wide variations in the degree of error between and within the models. Overall, each of the models, for a majority of the cases, were able to accuracy portray whether there was a negative or positive deviation from ideality, and in some cases, provided composition and pressure data that was accurate for the entire range with an average percent deviation of 0.75%. Due to the source of the calculated interaction parameters for each model being a function of the infinite dilution coefficient in two models, the majority of the errors can be seen in the middle of the composition range for the binary systems or areas where the model falsely predicts LLE. Although the method described in this paper does not always provide extremely accurate phase behavior values for composition and pressure, the method is quick and efficient for gaining a great amount of thermodynamic insight for a large quantity of available compounds. Utilizing the calculated parameters from the σ-profiles provided by Virginia Tech, a relativity accurate infinite dilution coefficient for a system could be calculated and phase behavior can be predicted using simplistic models provided. Future work into the utilization of σ-profiles could be a correlation to calculate the infinite dilution coefficient based on DFT and fitting MOSCED parameters based on the calculated coefficient instead of determining the parameters first. The method could also be improved from by using a more complex regression function with the addition of more group specific parameters particularly amines and carboxylic acids. The goal of the method was to remain simple while providing a maximum amount of accurate thermodynamic insight into the mixture, so more complex equations and a greater number of parameters were avoided. Finally, the use of other thermodynamic models to characterize the phase behavior could be used to accurately portray the systems. Again, the method was limited to quick calculations that could be performed in lieu of charts or graphs for solvent interactions, so the models with two interaction parameters where most useful. More complex models would, however, improve the accuracy of the calculated values for a predicted system.

5. References

1. Frank, T., Anderson, J., Olson, J., **2007**. Application of MOSCED and UNIFAC to Screen Hydrophobic Solvents for Extraction of Hydrogen-Bonding Organics from Aqueous Solution. Industrial & Engineering Chemistry Research 46, 4621–4625.
2. Elliott, J., Lira, C., **2012**. Introductory Chemical Engineering Thermodynamics, 2nd ed. Pearson Education, Inc.
3. Elliott, J., **2010**. A Simple Explanation of Complexation.
4. Lazzaroni, M., Bush, D., Eckert, C., Frank, T., Gupta, S., Olson, J., **2005**. Revision of MOSCED Parameters and Extension to Solid Solubility Calculations. Industrial & Engineering Chemistry Research 44, 4075–4083.
5. Mullins, E., Oldland, R., Liu, Y., Wang, S., Sandler, S., Chen, C.-C., Zwolak, M., Seavey, K., **2006**. Sigma-Profile Database for Using COSMO-Based Thermodynamic Methods. Industrial & Engineering Chemistry Research 45, 4389–4415.
6. Klamt, A., Eckert, F., **2000**. COSMO-RS: a novel and efficient method for the a priori prediction of thermophysical data of liquids. Fluid Phase Equilibria 172, 43–72.
7. Klamt, A., Jonas, V., Burger, T., Lohrenz, J., **1998**. Refinement and Parametrization of COSMO-RS. Journal of Physical Chemistry 102, 5074–5085.
8. Klamt, A., Eckert, F., **2004**. Prediction of vapor liquid equilibria using COSMOtherm. Fluid Phase Equilibria 217, 53–57.
9. Lin, S.-T., Sandler, S., **2002**. A Priori Phase Equilibrium Prediction from a Segment Contribution Solvation Model. Industrial & Engineering Chemistry Research 41, 899–913.
10. Danner, R., Gess, M., **1990**. A data base standard for the evaluation of vapor-liquid-equilibrium models. Fluid Phase Equilibria 56, 285–301.

6. **Appendix A: Infinite Dilution Coefficients**

7. **Appendix B: Calculated MOSCED Parameters**

8. **Appendix C: VLE Data**

9. **Appendix D: Calculated Interaction Parameters**